## REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER: NPRDC TR 84-31 | 2. GOVT ACCESSION NO. |
| 3. RECIPIENT'S CATALOG NUMBER | |
| 4. TITLE (and Subtitle): OFFICER CAREER DEVELOPMENT: DESCRIPTION OF AVIATION ASSIGNMENT DECISIONS IN THE ANTISUBMARINE WARFARE (ASW) PATROL COMMUNITY | 5. TYPE OF REPORT & PERIOD COVERED: Tech. Report, Oct 1980-Jan 1982 |
| | 6. PERFORMING ORG. REPORT NUMBER: 15-83-8 |
| 7. AUTHOR(s): Robert F. Morrison, Catarina Martinez, F. Wayne Townsend | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS: Navy Personnel Research and Development Center, San Diego, California 92152 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62763N, RF63-521-804-031-03.04 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS: Navy Personnel Research and Development Center, San Diego, California 92152 | 12. REPORT DATE: March 1984 |
| | 13. NUMBER OF PAGES: 31 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | 15. SECURITY CLASS. (of this report): UNCLASSIFIED |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Policy capturing
Decision theory
Officer quality index
Career development
Performance evaluation
Assignment decision

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

There is a shortage of senior, quality officers who are experts in specific functional areas such as tactics and those necessary to command major shore activities. A statistical policy-capturing technique was used to measure officer quality from detailer ratings of 134 male officers. Officer data cards and performance evaluation summaries were used as sources for the decision inputs. Analyses supported the hypotheses that (1) a simple decision theory is more effective with real-world data than is a complex theory, (2)

detailer ratings are consistent with Navy policy regarding operational (sea duty) performance, size of officer's peer group, and performance trend within a tour, (3) higher-rated officers obtain a graduate degree and a proven subspecialty in the same proportion as do lower-rated officers, and (4) detailer ratings are more highly correlated with performance evaluation data that compares an officer with his peers than it is with normative data that describes levels of performance. Findings are discussed in terms of their implications for (1) future research, (2) rater use of comparative vs. normative information, and (3) need for review of Navy policies to determine if they are consistent with the intent of policy makers.

NPRDC TR 84-31

MARCH 1984

# OFFICER CAREER DEVELOPMENT: DESCRIPTION OF AVIATION ASSIGNMENT DECISIONS IN THE ANTISUBMARINE WARFARE (ASW) PATROL COMMUNITY



NAVY PERSONNEL RESEARCH
AND
DEVELOPMENT CENTER
San Diego, California 92152

NPRDC TR 84-31 March 1984

# OFFICER CAREER DEVELOPMENT: DESCRIPTION OF AVIATION ASSIGNMENT DECISIONS IN THE ANTISUBMARINE WARFARE (ASW) PATROL COMMUNITY

Robert F. Morrison
Catarina Martinez
F. Wayne Townsend

Reviewed by
Edwin G. Aiken

Approved by
Martin F. Wiskoff

Released by
J. W. Renard
Captain, U.S. Navy
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

## FOREWORD

This research and development was conducted within exploratory development project RF63-521-804 (Manpower and Personnel Technology), work unit 031-03.04 (Personnel Distribution and Career Development). The purpose of the work unit is to determine factors within officers' careers that may be related to improvements in performance, the fulfillment of requirements for specific skills in senior ranks, and improvement in quality officer continuance.

This report is the fourth in a series being issued under this work unit. Previous reports described the factors that influenced the early career development of surface warfare officers (SWOs), background and initial sea tour factors that predicted SWO continuance beyond obligated service, and a pilot investigation of SWO career experiences and concerns (NPRDC TRs 82-59 and 83-6 and TN 83-11). The current report describes factors currently being used by detailers to assign officers to department head billets in the antisubmarine warfare (ASW) patrol community. It also assesses the relationship between those factors and several Navy policies and career opportunities. It is proposed that these factors are generalizable and can be used for research purposes across the air warfare communities.

Appreciation is expressed to CAPT Dallas Boggs (formerly NMPC-432) and LCDR Denzell Theis (formerly NMPC-432P) for their critical support and assistance in providing the data used in the research.

J. W. RENARD
Captain, U.S. Navy
Commanding Officer

J. W. TWEEDDALE
Technical Director

## SUMMARY

### Problem

While retention statistics for officers have been improving, there is a shortage of senior, quality officers who are experts in specific functional areas such as tactics and those necessary to command major shore activities. Research was initiated in FY81 to develop the technology that would aid in the solution to such problems.

### Objectives

The objectives of this research were to:

1. Develop a measure of officer quality within a single aviation community that could be used within the overall research work unit.

2. Assess the relationship between assignment decisions and Navy policy in the single aviation community.

3. Assess the relationship between subspecialties and graduate degrees acquired by officers and officer quality.

4. Assess the relationship between detailer decisions and use of comparative evaluation data versus normative-evaluation data.

### Method

A statistical policy-capturing approach was used to measure officer quality from detailer ratings of 134 male officers approaching reassignment in a single aviation community. Officer data cards and performance evaluation summaries were used as sources for the decision inputs. It was hypothesized that:

1. A simple decision theory is more effective with real-world data than are complex theories.

2. Detailer ratings are consistent with Navy policy.

3. A subspecialty or graduate degree is not related to higher quality ratings.

4. Detailers give more weight to supervisory data in which officers are rated against their peers than they do to data in which the officers are rated against a normative standard.

### Results

Analyses of three decision models, policies, subspecialty and graduate degree opportunities, and the use of comparative versus normative data showed that in the VP community:

1. A simple, three-variable, linear-regression model is more effective with real-world data than is a complex model.

2. Detailer ratings were correlated with policies that involved operational (sea duty) performance, size of officer's peer group, and performance trend within a tour.

3. Higher-rated officers obtained a graduate degree and a proven subspecialty in the same proportion as did lower-rated officers.

4. Detailer ratings were more highly correlated with performance evaluation data that compared an officer with his peers than they were with normative data that described an officer by his level of performance.

Conclusions

Results indicate that:

1. A simple, three-variable, linear-regression, decision-making model can effectively be used as an index of officer quality in future research.

2. Assignment decisions appear to reward officers contingent upon performance that is consistent with Navy policy.

3. The attainment of either a graduate degree or a proven subspecialty does not correlate with detailer decisions about quality. Therefore, it appears that higher-rated officers from the ASW patrol community do not find it attractive to obtain either a graduate degree or a proven subspecialty any more than lower-rated officers.

4. Performance evaluation data that compare an officer with his or her peers appear to provide clearer information for use in assignment decision making than do normative performance ratings.

Recommendations

It is recommended that:

1. Research personnel use the simple, three-variable, linear-regression, decision-making model as an index of officer quality in future research.

2. Policy makers retain in any future modifications to the "Report on the Fitness of Officers" the current emphasis that requires the rater to compare the officer ratee with his or her peers.

3. Research and aviation community personnel review early assignment policy regarding subspecialty tour experience and/or attendance at Naval Postgraduate School to clarify why officers with high ratings of their potential/promotability do not obtain a postgraduate education or a proven subspecialty more frequently than do officers with lower ratings.

4. Aviation community personnel review whether or not to continue policies that (a) give more weight to performance than other personal factors, (b) rate operational performance higher than shore duty, and (c) give an advantage to continued improvement during a single assignment.

## CONTENTS



## LIST OF TABLES

# INTRODUCTION

## Problem

Past research has concentrated on the retention of junior officers with little consideration toward qualities that contribute to efficient, senior line officers (Robertson & Pass, 1979; Githens, 1979; Cook & Morrison, 1982). The current emphasis on specialization in the surface warfare community (Walters, 1982), improvement in tactical competence for unrestricted line (URL) officers, increased numbers of senior line officers with the technical knowledge to manage major shore commands, and altered aircraft carrier (CV) command information center (CIC) billet structure[1] indicates a need to establish preliminary officer quality factors. Before the shortage of senior line officers can be addressed, it must first be determined what factors decision makers use to compare and judge officer quality.

Decision makers[2] are using large quantities and varieties of information to make selections, promotions, and assignments. The kinds of comparisons these decision makers make, the criteria they use for judgment, and whether the decisions are valid are important research problems because such decisions impact on the effectiveness of the organization as well as on the career of individuals who work within the organization. Although all such decisions are important, this research covers only promotion and assignment decisions.

Up to this point, promotion decisions have been neglected as a research topic. Stumpf and London (1981) have suggested that an investigation into the specific criteria used for management promotions needs to be undertaken and that this investigation should take place in an organizational setting. This research particularly fills the need expressed by Stumpf and London. In addition, it compares the applicability of various decision theories and investigates whether decision makers are able to implement the organization's statement of formal policy in their decisions.

It is concluded that, if senior line officers who possess the abilities necessary to manage major shore commands are to be available when needed, there must be a determination of (1) the criteria that decision makers use to make assignment decisions, which lead to promotion, (2) whether or not the decision criteria are consistent with Navy policy, and (3) a decision model that can be used as an index of officer quality for future research in the personnel distribution and career development project.

## Objectives

The objectives of this research were to:

1. Develop a measure of officer quality within a single aviation community that could be used within the overall research work unit.

2. Assess the relationship between assignment decisions and Navy policy in the single aviation community.

---

[1]COMNAVAIRLANT memorandum of 2 November 1982; subj: CV CIC officer billet upgrade.

[2]The term "decision maker" is used in a generic sense and refers to persons who contribute to the promotion and assignment of antisubmarine warfare (ASW) senior line officers.

3. Assess the relationship between subspecialties and graduate degrees acquired by officers and officer quality.

4. Assess the relationship between detailer decisions and use of comparative evaluation data versus normative evaluation data.

## Background

Several of the major theories of judgment and decision making were relevant to this research. These, along with the policy-capturing technique used to measure the types of information used and kinds of decisions made by individual decision makers, are described below.

1. Individual versus group decision making. Although some of these theories deal with individual decision making and others concentrate on group processes, and still others on social aspects, only the individual decision-making processes were relevant in this research.

2. High- versus low-risk decisions. Whether a decision was risky or riskless has been considered to be significant in the the decision-making process (Fischer, 1972). A decision was risky if it involved a number of alternatives and the decision maker did not know the consequences of each course of action but could express this uncertainty in terms of probabilities. A decision was riskless if the decision maker could specify, with certainty, the result associated with each alternative. Since the immediate consequences of the decisions involved in this research were known to the decision makers, they were termed riskless. Future decision consequences such as selection for promotion or command were not known, but other events and decisions intervened to reduce the risk of those consequences on the present assignment. The risks involved with the decisions in this study were minimal in the short term and therefore the decisions were assumed to be riskless.

3. Normative versus descriptive decision making. Decision theories have been categorized into normative and descriptive ones. Normative models prescribe how a person should make optimal decisions while descriptive models describe how a decision maker actually makes decisions.

a. Normative theories. Normative theories require or impose a validation of the decision and specify what a decision maker should do in order to be consistent and rational. Since this research did not address either the rationality of the decision maker or the validity of the decisions, only the descriptive aspects of decision making were relevant. The quality factor developed was based upon information cues as perceived by the decision makers (organization).

b. Descriptive theories. Descriptive theories comprise a set of statements that describe what a decision maker does. These can be used to predict the decision a person will make. There are several different types of descriptive models (Anderson, Deane, Hammond, & McClelland, 1981), three of which are discussed below.

(1) Process-tracing model. The process-tracing model presents a sequence of operations or thoughts (input) that is used to produce the resultant decision (output). Such a model is commonly developed by the researcher asking decision makers to describe their thinking (verbal protocol), observing the decision makers' eye movements, or by observing the decision makers' pattern of search for information (explicit information search) while making decisions (Payne, 1976). The first and last techniques were used in this research.

(2) Paramorphic model. Paramorphic models statistically describe input and output relationships but are not concerned with the intervening process. They are obtained through the use of a technique termed policy-capturing, which uses statistical analysis of the judgments to model (referred to as a policy) the process that generated those judgments. The most commonly used statistics in policy capturing are multiple regression, analysis of variance (ANOVA), conjoint measurement, and Bayesian probabilities. Traditional economists view the decision maker as one who makes continuous tradeoffs between value attributes, with the implication that the psychological process is compensatory in nature; that is, an increase in value on one attribute can compensate for a decrease in value on another. These compensatory models are based on additivity. Various studies have shown that the additive compensatory models can represent the judgment process very well (Tversky, 1967; Anderson, 1970; Sidowski & Anderson, 1967). The additive model has two derivatives, one based on the addition of stimuli to form an impression (summation); and the other, on the mean of the information (averaging) (Slovic & Lichtenstein, 1971).

In the compensatory approach, ANOVA may be used as a scaling procedure to obtain the desired additive representation (Fischer, 1972). However, a major drawback with an ANOVA design is that, in order to distinguish a summation process from an averaging process in the additive model, careful attention must be given to subtleties of stimulus construction and experimental design.

Other investigators have tried to bypass the difficulty of the ANOVA design with the use of the multiple regression paradigm. Such regression studies have restricted themselves to the assumption that the dimensions contribute linearly to overall value. Although these additive models are highly predictive, other investigators have argued that these decision makers actually utilized noncompensatory models.

(3) Noncompensatory model. Noncompensatory models combine attributes of both the process-tracing and paramorphic approaches. Glueck (1974) and Hogarth (1980) described four noncompensatory models: conjunctive, disjunctive, lexicographic, and elimination-by-aspects.

(a) The conjunctive or satisficing model is one in which the decision maker establishes a minimum acceptable level with respect to each attribute. Any alternative that falls below this level on any single attribute is considered unacceptable. Given a set of alternatives, a satisficing rule partitions this set into two subsets: those that are acceptable and those that are unacceptable. If, however, two or more alternatives are acceptable, additional consideration is introduced in order to choose between them. An example of a conjunctive model would be a job selection situation in which three criteria were used. If a candidate failed to meet the minimum standard for any of the three criteria, that individual would be automatically disqualified from any further consideration.

(b) A disjunctive model, on the other hand, is concerned with excellence rather than simple acceptability. In this model, a decision maker evaluates each alternative by its most outstanding attribute and the alterative with the most desirable attribute is selected. A decision maker disregards a low score on one attribute provided there is a very high score on one of the other dimensions. In keeping with the job selection example, a candidate who might be high in job aptitude (average for motivation), but low in intelligence, would be favored over a job candidate who might be average in all three attributes.

(c) The lexicographic model is one in which all the alternatives are compared with respect to the single important value attribute as established by the decision maker. If two or more alternatives are equivalent at this point, then the second most important attribute is taken into consideration. If comparison is insufficient for a decision to be made, then a third criterion is used, and so on. Differences on more important attributes cannot be compensated for by differences on less important attributes. In the job selection example, suppose that job aptitude was the most important variable and the two top candidates were equal on that measurement. In order to distinguish between them, the decision maker would have to look at the second most important variable. If candidates were equal on that variable, then the third variable would be considered, and so on, until a decisison could be made.

(d) The elimination-by-aspects model was developed by Tversky (1967) and is similar to the lexicographic model in that alternatives are eliminated by evaluating them on a sequence of attributes. However, this model proposes that attributes are selected in sequence based on a probabilistic scheme at each state in the decision process. All alternatives are evaluated on a single attribute and those that do not include the selected aspect are eliminated. Then, a second attribute is selected, the remaining alternatives are evaluated, and some are eliminated. The process continues until all the alternatives but one are eliminated. A major feature of the elimination-by-aspects model is that the probability of an alternative being selected depends on its overall value as well as on its relationship to the other available alternatives.

Although some may argue that the four preceding noncompensatory models are normative, they are typically classified as descriptive since they are not prescribing that decisions should be made according to these models nor do they assume rationality of the decision. However, these strategies do involve a quasi-logic in that some ordering of dimensions is involved and these methods seem intuitively justifiable. Nonetheless, there are severe biases associated with these strategies. For example, in the elimination-by-aspects formulation model, alternatives may be eliminated before consideration of their merits.

4. Policy-capturing technique.

a. Method. The policy-capturing technique describes information used by individuals to make decisions. This statistical method is used as an aid to obtain an objective description of the rules a decision maker uses to identify, select, and combine information as applied in the decision-making process. When the policy capturing approach is used to test a decision theory, the results describe the information processed by the decision maker according to the following parameters:

(1) Level of contribution made by each information element.

(2) Lack of information.

(3) Relationship among information elements.

(4) Representativeness of the information.

Although it is possible for a decision maker to have an almost infinite amount of information available, in most cases, the availability of information and its sources are finite. The policy-capturing technique requires that the available information be used to construct a data set to create information input variables called cues or predictors. The method also requires the decision maker to record final decisions about an array of stimulus objects in an unequivocal manner in order to establish a criterion

variable. Such decisions are normally based on a single criterion such as the job performance of individuals or the attractiveness of various job opportunities.

The stimulus scores are then subjected to statistical analysis that determines the weights of the information cues according to the influence that the cues have on the decisions. The model requires a high degree of intra- and inter-decision maker agreement.

b. Approach. The various statistical methods employed in policy-capturing research are classified into the Bayesian and the regression paradigms (Slovic & Lichtenstein, 1971). Both approaches have advantages and disadvantages concerning the statistical model being used and the type of information processed within them, as discussed below.

(1) Bayesian paradigm. The Bayesian paradigm is an iterative approach based on the principle that judgments are described initially by subjective or personal probabilities and revised optimally in accordance with Bayes' theorem as new information is acquired. A Bayesian analysis is derived from a distribution of probabilities based on what the results should be rather than on a set of actual observations. For this reason, the Baysian paradigm prescribes ways in which a decision maker can reach an optimal decision; that is, it provides a normative model. Because the present authors' purpose was to describe facts of the moment rather than to establish what should be the best combination of information to use to make later decisions, the Bayesian approach was not appropriate for the research herein.

(2) Regression paradigm. The regression paradigm is descriptive in nature and is divided into the ANOVA design and the correlation design, both of which are discussed below.

(a) ANOVA design. The ANOVA design can potentially detect linear, curvilinear, and configural aspects of the judgment process. Significant main effects for a cue would be that a decision maker's responses systematically vary with that cue while other cues are held constant. This would allow for trend analyses of main effects to be performed to detect linear, quadratic, cubic, or other curvilinear effects. Significant interaction between cues would imply that the effect on one cue would vary as a function of another, which would then produce a curvilinear or configural effect. A major drawback to the ANOVA approach is that less information is extracted than can be provided by the correlation paradigm, which directly weights the cues according to their contribution to the decision.

(b) Correlation design. The correlation design is based on linear regression. The description of this design by Slovic and Lichtenstein (1971) corresponds to the description of policy capturing in the paramorphic model. The focus of the correlation design is to model a judge's policy.

c. Measurement issues. Several measurement issues were relevant to this study as discussed below.

(1) Difference between ANOVA and correlation designs. According to Cohen (1968), the ANOVA and correlation designs are almost identical in their theoretical systems, although in actual practice differences arise. These differences generally favor the correlation design. Whereas the ANOVA design requires the data to be orthogonal, the correlation design takes into account the multicollinearity of the data and incorporates it into the statistical procedure. In the applied setting, orthogonal information is a rarity, if it occurs at all.

(2) Configural relationships and individual decision policy. A controversy exists as to whether configural relationships should be employed in the investigation of individual decision policy. It has been argued that linear models are robust enough to account for most of the variance (Dawes & Corrigan, 1974). Cohen (1968) found that nonlinear relationships did not add a significant amount of information and further may cause some information to be lost due to high multicollinearily.

(3) Rejection of judgment analysis (JAN). Judgment analysis or JAN is a design that has been used by multiple judges to study different policies (Naylor & Wherry, 1965). However, since Holzbach (1979) and Morrison[3] found that there was a very high level of interrater agreement among decision makers who faced three independent sets of decisions about the same task and information as used in this report, the JAN design was not used in this research effort. Here, the concentration was on documentation of information elements that decision makers used to make decisions and the relative weight given to each element.

(4) Real-time operation versus hypothetical data. Another aspect of the present study concerned the situation that previous researchers used in their research. Of eleven studies reviewed, only two used real-time operational data and decisions (Christal, 1967; Taylor & Wilstead, 1974), while the remaining studies used hypothetical data in a simulated situation (Brady & Rappoport, 1973; Hoffman, Slovic, & Rorer, 1968; Stumpf & London, 1981; Wherry & Naylor, 1966; Zimmer, 1981). The term "real-time" as used in this study referred to information that was derived from and decisions that were implemented in an actual organization. In addition, the consequences associated with decisions made were known to the decision maker and to the organization.

A major problem with the use of hypothetical data is that information is constructed to exclude multicollinearity of the variables. For example, Stumpf and London (1981) specifically designed their study to exclude any relationship among the variables by using a randomized block factorial design. Zedeck and Kafry (1977) also reported that the correlations among the dimensions used in their study approximated zero. One of the main reasons cited for use of a real-time approach was that the contribution of the elements were easier to interpret. Since a change in the correlation between any two variables will generally affect the whole multiple regression model (Darlington, 1968), and much of the data in any environment is correlated, an assumption of no multicollinearity is unrealistic (Zedeck & Kafry, 1977).

5. Relationships between decision making and organizational policy. The need to investigate whether the organization's statements of formal policy are related to decisions and might, therefore, be assumed to be implemented by the decision makers in their decisions was an important issue in the present promotion and assignment decision research. Past research showed that decision makers do not typically make their decision rules explicit (Schmidt & Kaplan, 1971). As a result, it was difficult to compare stated policy from an operational setting with the empirical policy derived via quantitative methods (Slovic & Lichtenstein, 1971; Stumpf & London, 1981). However, in the present research, it was possible to make this comparison.

---

[3]Morrison, R. F. Detailer inter-rater agreement (unpublished letter report). San Diego: Navy Personnel Research and Development Center, 20 November 1982.

6. Input information variables. The research reported here specifically defined personnel and performance information variables as input used by decision makers to make promotion and assignment decisions. Decisions were the product of (a) comparative and (b) normative information. Comparative information ranked candidate performance with other candidates. Normative information judged candidate performance according to standard rating scales. Morrison observed that in certain settings primary use was made of comparative data in personnel decisions and normative data received limited consideration (Sun's Computer, 1969).

## METHOD

### Subjects

Subjects were 134 male Navy officers from the ASW patrol community. All officers were scheduled for reassignment within a 9-month period and had the same approximate biography, as shown in Table 1. To predict quality ratings from given information variables, subjects were divided into two groups by assignment period, as shown in Table 2.

Table 1

Subject Biography

| Items | Attributes |
|---|---|
| Rank | LCDR |
| Age | Same |
| Qualifications | NFO or pilot |
| Years of service | 11 to 12 |
| Reassignment period | 9 months |

Table 2

Subject Groupings

| Variable | Group A | Group B |
|---|---|---|
| Population size | 62 | 72 |
| Assignment period | 1 July to 30 September 1981 | 1 October 1981 to 30 March 1982 |

## Variables

Independent and dependent variables were established. Independent variables were constructed from (1) personnel records, taken from individual officer data cards, and (2) performance summaries, taken from individual fitness reports, as shown in Table 3. Direct measures were made of work history, education, qualifications, comparative performance data, ratings of traits, estimates of promotability, and so forth. The direct measures, in turn, were used to construct other variables (e.g., average performance rating and performance trend during an assignment) that were used for analyses of hypotheses. Prior to any analysis, each variable was recoded so that the distribution of responses approximated a normal curve.

The dependent variable was a rating of the officer's overall quality that was made and used by detailers (the decision makers) to assign the officer to a squadron department head billet. Detailers assigned officers so that each squadron was equal in the quality of assigned officers. This decision criterion made each squadron relatively equal in its available managerial skill and ensured that several high potential officers would not be in direct competition with each other for available department head positions and on comparative performance evaluations.

Groups A and B were independently rated on a scale from one to ten (incremented in half-points) according to their potential for promotion and selection to command.[4] This measure of quality was the average rating given each officer by a group of three detailers. Previous researchers (Holzbach, 1979; Morrison, Note 3) have shown that inter-rater agreement among detailers involved in this task is very high ($r_{tt}$ = .95, .95, and .94 respectively).

## Hypotheses

It was hypothesized that (1) a compensatory, paramorphic model would account for more variance than two more complex models tested, (2) organizational policy statements would be correlated with the decisions, (3) achievement of secondary subspecialty requirements would not be related to the ratings, and (4) predictions from the use of regression equations with comparative data would be better than predictions from regression equations with noncomparative data.

## Procedure

Four strategies were used to test the four hypotheses.

1. The first strategy tested three decision models to determine whether a compensatory paramorphic model would account for the most variance. These models are described below as hypothesis 1a through 1c.

a. Hypothesis 1a tested a compensatory, paramorphic model to determine whether it would account for more variance than the other models tested (Dawes & Corrigan, 1974). This model assumed only an informal test of additivity as an underlying decision-making process.

---

[4]Personal communication with D. Theis, 12 March 1981.

Table 3

Independent Variables

| Information Variables | Description |
|---|---|
| **Personnel Record (officer data card)** | |
| WSKILL | One of two warfare skill classes (NFO or pilot). |
| EDUC | Graduate degree or no graduate degree. |
| ERLYPRM | Promoted more rapidly than norm in past. |
| SUBSPEC | A proven skill other than warfare skills; for example, carrier landing signal officer. |
| AQUAL | Acquisition of warfare skills beyond those required. |
| PERECORD | Sum of personal record variables. |
| **Performance Summary (fitness report)** | |
| Comparative Information[a] (Statistical Rating) | |
| PERF | Officer's overall rating on a 7-point statistical scale (i.e., "in the top 1%, in the top 5%"). |
| -SEA | Same rating but for sea tours only. |
| -SHORE | Same rating but for shore tours only. |
| AVELO | Average number of officers rated lower. |
| -SEA | Ratings for sea tours only. |
| -SHORE | Ratings for shore tours only. |
| Comparative Information (Peer-group Ranking) | |
| PROM | Average position among peers recommended for early promotion. |
| -SEA | Same but for sea tours only. |
| -SHORE | Same but for shore tours only. |
| TREND | Trend of performance evaluations during a tour. |
| -SEA | Trends for sea tours only. |
| -SHORE | Trends for shore tours only. |
| Normative Information[b] (Standardized Rating) | |
| SPEC | Average rating of 11 specific aspects of performance each rating on a 9-point scale. |
| -SEA | Ratings for sea tours only. |
| -SHORE | Ratings for shore tours only. |
| TRAIT | Average of ratings of six personal traits (9-point scale). |
| -SEA | Ratings for sea tours only. |
| -SHORE | Ratings for shore tours only. |
| POTENT | Average rating of potential to work in five types (9-point scale). |
| -SEA | Ratings on sea tours only. |
| -SHORE | Ratings for shore tours only. |
| PERFSUM | Sum of all performance ratings, PERF, AVELO, TREND. |
| SIZE | Average number of peers in the officers' comparison group (comparison group size). |

[a]Comparative information ranked candidate performance with other candidates.

[b]Normative information judged candidate performance according to standard rating scales.

b. Hypothesis 1b tested a two-step mixed process tracing model developed with the use of the verbal protocol technique (VPT). It was conjectured that the detailers may have been trying to emulate the formal selection board (panel of superior officers) that would review the subject officers in approximately 2 years for promotion to CDR. The selection board in question makes its decisions in a two-step process. On the first quick pass, the most qualified candidates are immediately recommended for promotion and the least qualified are immediately excluded from further consideration. On the second pass, the middle group is considered very thoroughly, and the more highly qualified of this restricted group are recommended for promotion. For this research, it was assumed that the variables used in each step were different. Such an assumption was necessary to accommodate differences that cannot be represented by a simple linear regression equation.

The model was assumed to be noncompensatory (quasi-conjunctive) in step one because those with a score above a high standard on one or more variables were considered promotable and those with a score below a low standard on the same variable(s) were excluded from further consideration. The middle group, with scores between the high and low standards on the same variable(s), was submitted to step two. The model was assumed to be compensatory in step two. It was hypothesized that predictions from the use of this model would not be significantly different than predictions from the use of the compensatory paramorphic model analyzed first.

c. Hypothesis 1c tested a singular process-tracing model developed from the explicit information search technique (EIS). It was assumed that the decision makers (1) scanned the entire set of information available for an individual, (2) grouped those elements that were perceived as similar into several clusters, (3) combined the data within each cluster in an additive manner to yield a cluster score, (4) weighted each cluster score according to a pre-established set of decision rules, and (5) aggregated the cluster scores additively to obtain an overall score. It was hypothesized that predictions from the use of this model would not be significantly different than predictions from the use of the compensatory paramorphic model.

2. The second strategy used four analyses to test whether detailers' ratings were related to explicit organizational policy statements, as described below in hypotheses 2a through 2d. All analyses used a form of the verbal protocol technique from the process-tracing model. The Navy's explicit policies regarding which officers would be selected for key assignments and promotion were collected by researchers (a) reviewing policy manuals, (b) listening to presentations made to top management, (c) observing interviews between detailers and officers of several ranks, and (d) interviewing detailers. The policy statements were then converted into the following specific hypotheses regarding the relationship between the information and the detailers' ratings:

a. Hypothesis 2a. Personal history data carry less weight when making decisions than do performance data.

b. Hypothesis 2b. Fitness reports that describe performance during operational line assignments (sea tours) are given more weight than nonoperational line assignments (shore tours) fitness reports.

c. Hypothesis 2c. The trend of performance during a tour of duty (i.e., going from relatively lower to higher) will affect ratings in a positive way.

d. Hypothesis 2d. The size of the comparison group in which an officer is rated will affect ratings, with larger comparison groups providing a positive bias to the overall rating.

3. The third strategy tested the hypothesis that the obtainment of a master's degree and/or proven subspecialty beyond the warfare skill would not be correlated with the detailers' ratings (estimates of promotability to CDR). Although the Navy considers it desirable for an officer to obtain a proven subspecialty in addition to the primary warfare specialty, there are disincentives for the officer to do so. Officers have a better chance of obtaining a proven subspecialty if they obtain a master's degree in a field and program that have been approved by the Navy. The major focus of such degrees is the Naval Postgraduate School (NPS). However, there are major disincentives for junior officers who are considering attending NPS; one is the additional service required to repay the Navy for its investment in them and a second is the imposition of an educational utilization assignment that takes the officers away from their primary warfare specialty and the officers who make up that warfare community.

4. The fourth strategy used two analyses to test the hypothesis that predictions made from regression equations with comparative data would be better than predictions made from regression equations with noncomparative data. These analyses tested the preestablished rule that information that compares an individual with peers is better than information that compares the individual to a normative standard.

## Analyses

Validation of models has been a neglected issue in the research of decision making (Zimmer, 1981); that is, in studies that constructed models, only one was found that tested the predictive validity of that model. The present research provided checks for all models developed.

One validation method described by Norman (1965) and adapted by Morrison and Hinrichs (1980) is the double cross-validation design. This design draws two independent samples from the population and develops regression models on each sample. A variable that is present in only one model is assumed to account for error variance associated with sample uniqueness and is excluded from the final model. The final model is then constructed on all cases; that is, it combines both samples. This method was used to test the decision-making hypotheses with the exception that the two "samples" were actually time-constrained populations; that is, all officers from that community had been reassigned during the specific time period.

Hypothesis 1a of the first strategy was tested by the double cross-validation model. All 29 variables were submitted to a stepwise multiple regression process. Each variable was submitted separately for each of groups A and B, as shown in Table 4. The analysis was halted when the increase in variance accounted for by an additional variable became nonsignificant. The final model was constructed from the total sample size (N = 134). Only variables that appeared in both groups A (N = 62) and B (N = 72) were used.

Hypothesis 1b of the first strategy (the two-step VTP, mixed-process-tracing model) was tested by a search of the independent variables to identify a high standard for use in rating prediction. The top 30, 35, or 40 percent of the sample size was searched and identified as high. The bottom 10, 15, or 20 percent of the sample size was searched to determine if one or more variables provided accurate selection for identification of a low group. The middle 50 percent was submitted to a stepwise multiple regression procedure to provide the final cut between the high and low groups.

Table 4

Correlations Between Information Variables and Potential Ratings

| Information Variables | Potential Ratings/Dependent Variables | | |
|---|---|---|---|
| | Group A (N=62) | Group B (N=72) | Total (N=134) |
| **Personnel Record (PERECORD) (obtained from officer data card)** | | | |
| WSKILL | -.21 | -.06 | -.12 |
| EDUC | -.06 | -.03 | -.04 |
| ERLYPROM | -.38** | -.25* | -.32** |
| SUBSPEC | .19 | .10 | .15 |
| AQUAL | .09 | .20 | .15 |
| PERECORD | -.18 | -.05 | -.12 |
| **Performance Summary (PERFSUM) (obtained from fitness report)** | | | |
| PERF | .74** | .71** | .72** |
| -SEA | .68** | .67** | .68** |
| -SHORE | .56** | .44** | .50** |
| AVELO | .67** | .66** | .66** |
| -SEA | .77** | .62** | .68** |
| -SHORE | .11 | .41** | .24** |
| PROM | .77** | .71** | .74** |
| -SEA | .82** | .71** | .76** |
| -SHORE | .60** | .66** | .63** |
| TREND | .71** | .73** | .72** |
| -SEA | .66** | .58** | .61** |
| -SHORE | .52** | .54** | .53** |
| SPEC | .65** | .57** | .60** |
| -SEA | .30** | .41** | .37** |
| -SHORE | .39** | .41** | .40** |
| TRAIT | .61** | .61** | .61** |
| -SEA | .37** | .43** | .40** |
| -SHORE | .00 | .23 | .13 |
| POTENT | .60** | .58** | .59** |
| -SEA | .38** | .45** | .42** |
| -SHORE | .44** | .56** | .51** |
| PERFSUM | .80** | .76** | .78** |
| SIZE | .03 | .00 | .02 |

*p < = .05.
**p < = .01.

Hypothesis 1c of the first strategy (the EIS, process-tracing model) was tested by the calculation of a factor analysis of the total set of independent and dependent variables. Factors were then entered into a hierarchical multiple regression procedure separately on each group according to the magnitude of the dependent variable's factor loading on the factor. Significant factor variables that were common to both groups were then submitted to a multiple regression procedure on the total sample.

To test all five hypotheses from the second and third strategies, product-moment correlations were computed among the independent variables and between the independent and dependent variables. Comparisons among the correlations were made with the assumption that the samples were not independent (Peatman, 1963).

To test strategy four for differences in the predictive ability of comparative and noncomparative information, multiple regression equations were developed. First, a multiple regression equation was developed on group B (N = 72) with the use of comparative information, and then one was developed on the same group with the use of noncomparative information. Three variables were used to measure noncomparative information and four variables were used to measure comparative information (see Table 3). To equalize the number of independent variables for each type, a stepwise multiple regression procedure was stopped when three variables had entered the equation for comparative information. The results of the two procedures were applied to group A (N = 62) to develop two sets of predicted scores. The correlations between the predicted scores for each model and the actual scores were then compared.

## RESULTS

The effects of four strategies used to analyze decision models, policy implementation, and comparative versus normative data were as follows.

### Strategy 1: Decision Models

Hypothesis 1a was accepted since the results were not only statistically significant but they were also operationally useful (66% of the variance in the ratings was accounted for) in the establishment of individual quality indices for research purposes. When the 29 variables were submitted to the multiple regression procedure, similar results were obtained in both groups. Both achieved a multiple correlation of .84 before additional variables failed to provide a significant increase in the multiple regression coefficient. These results were obtained with three variables in group A and four variables in group B. The final model obtained for the total sample achieved a multiple correlation of .81 with the three variables that were common to the equations for groups A and B ($p < .001$, df = 127). The three variables were: (1) the average standing on comparison data (PERF), (2) the ratio of times recommended for early promotion (PROM), and (3) the relative ranking among peers represented by the average number rated lower (AVELO).

Hypothesis 1b was rejected because explicit cut scores could not be identified to segment the total group in step one into high, middle, and low subgroups as required in the hypothesis. When cross-tabulations were conducted by a comparison of scores on each independent variable and the top 30, 35, and 40 percent on the dependent variable for group A and group B independently, no variable or combination of variables achieved the required separation between the high group and the rest. The same results occurred when cross-tabulations were conducted by a comparison of scores on each independent variable and the bottom 10, 15, and 20 percent of groups A and B on the dependent variable. The middle 50 percent of the groups could not be identified to continue the test of the VPT, mixed-process-tracing model.

Hypothesis 1c was not supported. The factor analytic approach did not attain results as good as the results from the regular stepwise procedure. Three factors were identified that had factor loadings of .30 or higher for the dependent variable, met the scree test (Tatsuoka, 1971), and accounted for 41 percent of the variance. The three factors were

represented by independent variables that described performance evaluations received during different assignments. The first factor summarized the most recent operational line assignment (sea tour); the second, the initial nonoperational assignment (shore tour); and the third, the current (second) shore tour. The second factor was dropped because of too much missing data--many officers' records did not include the initial shore tour in the 8-year-reporting period. Scores for the remaining two factors were computed with the addition of the scores for each independent variable that had a factor loading of .60 or higher. When factor scores were entered into a multiple regression procedure, 59 percent of the variance was accounted for ($R1.23 = .77$, $df = 129$, $p < .001$). However, this did not equal the multiple correlation coefficient of .81 that was found in the stepwise procedure.

Strategy 2: Policy Implementation

Hypothesis 2a was supported. The correlation between the sum of personnel record variables (PERECORD) and the detailer ratings was not significant (see Table 4). The correlation between the sum of the performance summary variables (PERFSUM) and the detailer ratings was significant ($p < .001$). The correlation between the two sums was -0.11. Thus, performance rating information contributed more to the formation of overall ratings than did personal history information.

Hyptohesis 2b was supported. The correlations between the dependent variable and sea tour performance evaluation variables were not always significantly greater than those between the dependent variable and shore tour performance evaluation variables (see Table 4). Correlations that involved sea tour variables were significantly greater than those that involved shore tour variables for four out of seven variables. Only one of the seven comparisons would have been significant by chance. Although the tests of the seven variables were not strictly independent, it was considered that these results were significant ($\chi^2(2) = 85.8$, $p < .001$).

Hypothesis 2c, which conjectured that a trend of improved performance during a tour would be a significant factor in the evaluation of officers, was supported. The correlation between the average trend of performance evaluations within assignments and the detailer ratings was significantly different from zero ($p < .001$) (see Table 4).

Hypothesis 2d was supported. The correlation between the average performance rating of the officer (PERF) and the detailers' rating was .73 ($p < .001$) (see Table 4). When the average size of the peer group was also included in the assessment of the data, the multiple regression coefficient obtained was .76. This increase was statistically significant ($F(1,131) = 21.83$, $p < .001$), which indicated that the detailers had considered both the overall performance of an officer and the size of the comparison group.

Strategy 3: A Master's Degree or Subspecialty

Hypothesis 3 was supported. As shown in Table 4, the correlations between the presence of either a proven subspecialty (SUBSPEC) or an advanced degree (EDUC) and the detailers' ratings do not achieve significance at the $p < .05$ level for either group A or B or for the total. The officers rated highest by the detailers did not choose to obtain a proven subspecialty or an advanced degree at a greater rate than did those rated lower.

Strategy 4: Comparative Versus Normative Data

Hypothesis 4, which proposed that comparative performance rating data would be a better predictor of detailer decisions than normative performance rating data, was

supported. When a linear model was formed on group B from the three best of the four variables (PERF, PROM, AVELO, & TREND) that compared an officer to his peers, the regression coefficient was .82. When this equation was applied to group A and the criteria were predicted, the correlation between the predicted detailer ratings and the actual detailer ratings ($r_{y_c^1 y}$) was .83.

The same type of analysis was performed using those variables derived from normative standards and yielded a regression coefficient of .64 on group B. When this equation was applied to group A and the criterion predicted, the correlation between this predicted detailer and the actual detailer rating did not change ($r_{y_n^1 y} = .64$).

To assess whether there was a significant difference between $r_{y_c^1 y}$ and $r_{y_n^1 y}$, the t-test for dependent correlations was employed (Cohen & Cohen, 1975). It was found that $r_{y_c^1 y_n^1} = 0.65$ and the resulting test of the difference between $r_{y_c^1 y}$ and $r_{y_n^1 y}$ was significant ($t(59) = 3.22$, $p < .01$).

## DISCUSSION AND CONCLUSIONS

### Policy Capturing

Consistent with previous research (Slovic, Fischoff, & Lichtenstein, 1977), the single linear model of information provides a better prediction of detailer decisions than the other methods used to combine data. Only three variables out of 29 were required to provide a reliable, effective means for the prediction of detailer ratings for research purposes. The simplicity of the model may be explained in more than one way. The decision makers may have established only a small number of variables that they considered useful. On the other hand, there may have been so much multicollinearity present in the available data that the detailers used large quantities of information as reliability checks against the variables they considered. It should be noted that 34 percent of the variance remained to be accounted for.

Although Darlington (1968) stipulated that multicollinearity contributes to the instability of Beta weights, the results of hypothesis 1a, which used the double, cross-validation design, do not indicate that this is a major problem in the research. Darlington's suggested use of factor analysis to compensate for high multicollinearity does not prove to be effective in this research, as shown in the test of hypothesis 1c. It appears that the position of Hobson, Mendal, and Gibson (1981) that artificially orthogonal groupings of stimuli may not validly describe the past decision making of raters is upheld. As evidenced by the above, the results obtained when raw data were used with simple linear regression are significantly different from results obtained when linear regression is applied to factor analytically restructured data. Also, the raw data results are more realistic and more likely to be useful to the decision makers in the organization.

As indicated by Hobson et al. (1981), the performance appraisal is typically used to accomplish two major functions: (1) organizational control and (2) individual development. The policy-capturing technique can provide a useful framework by which to describe and understand these two major functions. Since individual development is not a major factor in the decisions used in this research, the results are discussed only in terms of organizational control.

To assess the effectiveness of a performance appraisal system's contribution to organizational control, the relationship between specified, desired behavior and organizational rewards should be evaluated. In this research, organizational rewards are defined as higher ratings that may lead to a better assignment, which in turn may lead to promotion. The desired behaviors are explicitly defined as hypotheses 2a through 2d and were subjected to a policy-capturing technique.

The results from the tests of hypotheses 2a through 2d indicate that decision makers are able to take forms of officer behavior, expressed as clearly defined statements of policy, and implement them in operational decisions. The detailers clearly consider performance evaluation data to be more important than personnel file information. The type of assignment (sea vs. shore) in which an officer performs has some contribution to the decision but not to the extent that performance only in sea billets rather than performance in a mix of sea and shore billets increases the officer's opportunity to be selected for a top billet. It is clear that improving performance within any single assignment helps an officer obtain a high rating, and this effect is enhanced if the officer is judged in competition with a large group of peers. The achievement of organizational rewards appears to be contingent upon behavior that is consistent with expressed policy. Therefore, if the Navy desires to attract more top quality officers to the graduate education and proven subspecialty programs, it must establish some means by which to reward high quality officers who obtain a subspecialty prior to the 20-year mark at which time they are considered for promotion to captain.

While Stumpf and London (1981) found that different decision makers used different criteria in promotion decisions, those results were not replicated in this research. It appears that, when the decision process is formalized and decision makers are given specific information on organization policy, the information is integrated into consistent decisions as proposed by Stumpf and London. Additional support is provided by the high interrater agreement among various groups of decision makers who are faced with the same task at various periods of time (Holzbach, 1979; Morrison, Note 3).

A final supporting element for the position that organization policy is being implemented effectively is provided by previous research with junior surface warfare officers (Holzbach, 1979). Because the samples are different and the data are recoded to approximate a normal distribution in this reseach, the results from hypothesis 1a of this research cannot be compared statistically with Holzbach's work. However, similarities can be noted because nearly identical variables entered the multiple regression equations for each sample of officers. The prediction equation for the junior surface warfare officers is:

$$Ys = 9.44 - 1.22\,Xa + 1.58\,Xb - 0.54\,Xc,$$

where

Xa = average rating on the comparison data (one equals a high rating),

Xb = average number of times recommended for early promotion, and

Xc = relative standing among peers with the average number rated above one on the comparison data.

The prediction equation for the aviation officers is:

$$Ya = 2.33 - .28\,X1 + .33\,X2 + .26\,X3,$$

where

$Xa \approx X1$ = normalized average ratings on the comparison data,

$Xb \approx X2$ = normalized average position among peers who are recommended for early promotion, and

$-Xc \approx X3$ = normalized relative standing among peers with the average number rated below one on the comparison data.

It is apparent from the above results that the two major problems that interfere with the use of performance appraisals as a means of organizational control were not encountered in this study. Supervisors did not lack awareness of the content and the meaning of their appraisals nor did they tend to inconsistently and inefficiently combine information when they made appraisals in contrast to the expectations of Hobson et al. (1981). The clear definition of organizational policy appears to minimize the problems.

Decision-making Theory

To provide a comprehensive and reliable model of the detailers' decisions, various decision theories were used to develop the analytical procedures that were applied. Although this research is not a pure test of the theories, the results provide information that may be useful for subsequent research in the area of decision making and information processing.

Hypothesis 1a assumes a paramorphic compensatory model in the use of simple, three-variable, linear regression to capture the policy of three decision makers. Even though the use of a regression procedure does not provide a literal test of the additivity required in a compensatory model, the underlying process of the statistical procedure is assumed to be additive or compensatory in nature. Given this assumption, the model obtained from the results of hypothesis 1a show that the decision was somewhat compensatory by the fact that a low value on one of the variables in the model is compensated for by a high value on another variable.

Although the underlying process could have been assumed to be compensatory, the regression procedure itself is not a straightforward compensatory rule since the Beta weights are not unitary. Given this, a high value on one variable does not directly compensate for a low value on another variable. Implications for modification of the strict compensatory model are that (1) an additivity rule is employed when decisions are made and (2) variables of less importance do not outweigh variables of more importance in the decision process. If this is true, then the regression procedure can be seen as a combination of compensatory and noncompensatory rules.

Hogarth (1980) stated that the requirement that multiple regression attributes be independent serves as a major drawback to the use of the technique to model decision making. Although this may be a theoretical problem, it did not arise in this research as demonstrated in the results obtained via the double, cross-validated procedure. As a predictive model, the multiple regression model is remarkably accurate and consistent in its ability to predict judgments in both laboratory and applied settings. The consistency remained high, even when theoretically incompatible assumptions such as noncompensatory processes about decision making were made, and they accurately predicted judgments generated by other processes.

Hogarth (1980) has further stated that the strategies people use in decision making are affected by a number of factors such as (1) the number of alternatives, (2) the number of attributes per alternative, (3) the order of presentation, (4) their familiarity with the decision task, and so forth. Hogarth also stated that combinations of compensatory and noncompensatory strategies are sometimes used. Payne (1976) found that when subjects were faced with a two alternative situation, they employed strategies consistent with the compensatory model. However, when these same subjects were given a more complex task that consisted of more than two alternatives, they employed strategies consistent with the conjunctive or elimination-by-aspects models to reduce the amount of alternatives as quickly as possible.

Hypothesis 1b was formulated with the use of a multistep, mixed-process-tracing model that incorporates a form of the verbal protocol technique. The detailers described the behavior of promotion boards whose results they were attempting to predict. The detailers reported that two steps were employed by the board to cut down on the time required to make its promotion decisions.

The researchers assumed that the first step in the process was noncompensatory, as postulated in hypothesis 1b. When this hypotehsis was rejected, the possibility that the first step was compensatory was investigated with the use of separate stepwise multiple regressions to identify the high group (top 31%) versus the rest (69%) and the low (17%) versus the rest (83%). A single variable, PERF, became the one that formed the regression equations for both the high group and the low group. When the middle group (52%) was submitted to the stepwise multiple regression procedure, a second, single variable (PROM) accounted for a significant portion of the variance. When the predicted ratings obtained from this two-step, compensatory-process-tracing model were correlated with the actual ratings, significant results were obtained ($r = .32$, $p < .001$). However, the predicted ratings generated by the compensatory paramorphic model were correlated with the actual ratings ($r = 0.81$, $p < .001$) to a significantly higher degree ($z = 3.09$, $p < .001$) than the results from this mixed model. In future research, it appears that a compensatory model might be assumed to produce better results than a conjunctive noncompensatory model.

However, this is not conclusive evidence that a multistep mixed strategy was not used since the detailers faced a much less complex decision situation than the one present in a promotion board. This research did not formally test the various compensatory and noncompensatory models given the type of data used. Since the data used in this study were real, experimental manipulation was not possible. Therefore, many of the hypotheses are not clearly defined as specific compensatory or noncompensatory models. If the inability to clearly classify models is as typical of other situations as it is here, then the applicability of the results of laboratory research on descriptive models may be questionable. It may be necessary to design research so that descriptive models can be investigated in an applied setting with the use of real data.

Comparative Data

The results from strategy four indicate that decision makers find comparative information much more useful than normative information. This is consistent with Morrison's observation (Sun's Computer, 1969) that involved executive restaffing during a merger. These results also indicate that some of the current legislation, such as the Civil Service Reform Act of 1979, that prohibits the use of comparative information for performance evaluation purposes (Landy & Farr, 1980) may be detrimental to effective decision making.

The research shows that the use of the multiple regression procedure can be effectively applied to capture a decision maker's policy in an applied setting where there is high multicollinearily among the variables. It also shows that decision makers can incorporate specific statements of formal organizational policy into their decisions effectively and consistently. However, tests of descriptive decision theory are not supported by the results of this study. The test of comparative versus noncomparative performance evaluation information shows a significant advantage of comparative over normative information.

## RECOMMENDATIONS

This investigation only documents the inputs made by detailers to rate officers for future assignment and does not validate these criteria in relation to resultant decisions. Based on this limited consideration, it is recommended that:

1. Research personnel use the simple, three-variable, linear-regression, decision-making model as an index of officer quality for future research in the personnel distribution and career development project. The model represents typical ASW patrol community officers and may be generalized to other air warfare communities.

2. Aviation community personnel retain in any future modifications of the officer fitness report the current emphasis that compares an officer ratee with his or her peers.

3. Aviation community and research personnel review early assignment policy regarding subspecialty tour experience and/or attendance at Naval Postgraduate School to clarify why officers with high ratings have not obtained a postgraduate education or a subspecialty more frequently than have officers with lower ratings.

4. Aviation community personnel review whether or not to continue the policies that (a) give more weight to performance than other personal factors, (b) rate operational performance higher than shore duty, or (c) give an advantage to continued improvement during a single assignment.

5. Research personnel conduct decision-theory research in applied settings in which decision makers must actually integrate several pieces of information and produce a resultant decision.

## REFERENCES


Anderson, N. H. Functional measurement and psychophysical judgment. Psychological Review, 1970, 77, 153-170.

Anderson, B. F., Deane, D. H., Hammond, K. R., & McClelland. Concepts in judgment and decision research: Definitions, sources, interrelations, comments. New York: Praeger, 1981.

Brady, D., & Rappoport, L. Policy capturing in the field: The nuclear safeguards problem. Organizational Behavior and Human Performance, 1973, 9, 253-256.

Christal, R. E. Selecting a harem and other applications of the policy capturing model (PRL-TR-67-1). Lackland Air Force Base, TX: Personnel Research Laboratory, March 1967. (AD-658 025)

Christal, R. E. JAN: A technique for analyzing group judgment. Journal of Experimental Education, 1968, 36, 24-27.

Cohen, J. Multiple regression as a general data analytic system. Psychological Bulletin, 1968, 70, 426-443.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum, 1975.

Cook, T. M., & Morrison, R. F. Surface warfare junior officer retention: Early career development factors (NPRDC Tech. Rep. 82-59). San Diego: Navy Personnel Research and Development Center, August 1982. (AD-A118 717)

Cook, T. M., & Morrison, R. F. Surface warfare junior officer retention: Background and first sea tour factors as predictors of continuance beyond obligated service (NPRDC Tech. Rep. 83-6). San Diego: Navy Personnel Research and Development Center, January 1983. (AD-A124 325)

Darlington, R. B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.

Dawes, R. M., & Corrigan, B. Linear models in decision making. Psychological Bulletin, 1974, 81, 95-106.

Fischer, G. W. Multi-dimensional value assessment for decision making (Rep. No. 037230-2-1). Ann Arbor, MI: Office of Research Administration, June 1972. (AD-764 156)

Githens, W. H. Navy officer exit statement analysis (NPRDC Spec. Rep. 79-15). San Diego: Navy Personnel Research and Development Center, April 1979.

Glueck, W. F. Decision making: Organization choice. Personnel Psychology, 1974, 27, 77-93.

Hobson, C. J., Mendel, R. M., & Gibson, F. W. Clarifying performance appraisal criteria. Organizational Behavior and Human Performance, 1981, 28, 164-188.

Hoffman, P. J. Cue consistency and configurality in human judgment. In B. Kleinmuntz (Ed.), Formal Representation of Human Judgment. New York: Wiley, 1968, 53-90.

Hoffman, P. J., Slovic, P., & Rorer, L. G. An analysis of variance model for the assessment of configural cue utilization in clinical judgment. Psychological Bulletin, 1968, 69, 338-349.

Hogarth, R. M. Judgment and choice: The psychology of decision. New York: Wiley, 1980.

Holzbach, R. L. Surface warfare junior officer retention: Problem diagnosis and a strategy for action (NPRDC Tech. Rep. 79-29). San Diego: Navy Personnel Research and Development Center, August 1979. (AD-A073 463)

Landy, F. J., & Farr, J. L. Performance rating. Psychological Bulletin, 1980, 87, 72-107.

Madden, J. M. Policy-capturing model for analyzing individual and group judgment in job evaluation. Journal of Industrial Psychology, 1964, 2, 36-42.

Maguire, T. O., & Glass, G. V. Component profile analysis (COPAN): An alternative to PROF. Educational and Psychological Measurement, 1968, 28, 1021-1033.

Morrison, R. F. Officer career development: Surface warfare officer interviews (NPRDC TN 83-11). San Diego: Navy Personnel Research and Development Center, July 1983.

Morrison, R. F., & Hinrichs, J. R. R&D is different: Staffing laboratory suport activities. Paper presented at the meeting of the Annual Academy of Management, Detroit, August 1980.

Naylor, J. C., & Wherry, R. J., Sr. The use of simulated stimuli and the JAN technique to capture and cluster the policies of raters. Educational and Psychological Measurement, 1965, 25, 969-986.

Norman, W. T. Double-split cross-validation: An extension of Masier's design, two undesirable alternatives, and some enigmatic results. Journal of Applied Psychology, 1965, 49, 348-357.

Payne, J. W. Task complexity and contingent processing in decision making: An information search and protocol analysis. Organizational Behavior and Human Performance, 1976, 16, 366-387.

Peatman, J. G. Introduction to applied statistics. New York: Harper and Row, 1963.

Robertson, D., & Pass, J. J. Relationship of officer first assignment and educational major to retention (NPRDC Tech. Rep. 79-12). San Diego: Navy Personnel Research and Development Center, March 1979. (AD-A067 666)

Schmidt, F. L., & Kaplan, L. B. Composite vs. multiple criteria: A review and resolution of the controversy. Personnel Psychology, 1971, 24, 419-434.

Sidowski, J. B., & Anderson, N. H. Judgments of city-occupation combinations. Psychonomic Science, 1967, 7, 279-280.

Slovic, P. Analyzing the expert judge: A descriptive study of a stock broker's decision processes. Journal of Applied Psychology, 1969, 53, 255-263.

Slovic, P., Fischoff, B., & Lichtentein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.

Slovic, P., & Lichtenstein, S. Comparison of bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Stumpf, S. A., & London, M. Capturing rater policies in evaluating candidates for promotion. Academy of Management Journal, 1981, 24, 752-766.

Sun's computer helps find heir. Business Week, 4 October 1969, p. 83.

Taylor, R. L., & Wilsted, W. D. Capturing judgment policies: A field study of performance appraisal. Academy of Management Journal, 1974, 17, 440-449.

Tatsuoka, M. M. Multivariate Analysis. New York: Wiley, 1971.

Tversky, A. Addivity, utility, and subjective probability. Journal of Mathematical Psychology, 1967, 4, 175-202.

Walters, R. L. Personal from the Deputy Chief of Naval Operations (Surface Warfare). Surface Warfare, 1982, 7(8), inside front cover.

Wherry, R. J., Sr., & Naylor, J. C. Comparison of two approaches--JAN and PROF--for capturing rater strategies. Educational and Psychological Measurement, 1966, 26, 267-286.

Zedick, S., & Kafry, D. Capturing rater policies for processing evaluation data. Organizational Behavior and Human Performance, 1977, 18, 269-294.

Zimmer, I. A comparison of the prediction accuracy of loan officer and their linear additive models. Organizational Behavior and Human Performance, 1981, 27, 69-74.

## DISTRIBUTION LIST

Chief of Naval Operations (OP-115) (2), (OP-135C4), (OP-140F2), (OP-987H)
Deputy Chief of Naval Operations (Aviation), (OP-59), (Manpower, Personnel and Training) (OP-130E) (3), (OP-136D2)
Chief of Naval Material (NMAT 0722)
Chief of Naval Research (Code 270), (Code 440) (3), (Code 442), (Code 442PT)
Chief of Naval Education and Training (00A), (N-21)
Chief of Naval Technical Training (016)
Commandant of the Marine Corps (MPI-20)
Commander Naval Military Personnel Command (NMPC-013C), (NMPC-4), (NMPC-43) (5)
Commander, Navy Recruiting Command (Code 20)
Commanding Officer, Naval Aerospace Medical Institute (Library Code 12) (2)
Commanding Officer, Naval Technical Training Center, Corry Station (Code 101B)
Commanding Officer, Naval Training Equipment Center (Technical Library) (5), (Code N-1)
Commanding Officer, Office of Naval Research Branch Office, Chicago (Coordinator for Psychological Sciences)
Superintendent, Naval Postgraduate School
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria (PERI-ASL), (PERI-ZT), (PERI-SZ)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Manpower and Personnel Division), (Scientific and Technical Information Office)
Commander, Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/OT), (CNET Liaison Office AFHRL/OTLN)
Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (AFHRL/LR)
Commanding Officer, U.S. Coast Guard Research and Development Center, Avery Point
Institute for Defense Analyses, Science and Technology Division
Defense Technical Information Center (DDA) (12)

U211196

DEPARTMENT OF THE NAVY

NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA 92152

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

POSTAGE AND FEES PAID
DEPARTMENT OF THE NAVY
DOD-316